

⑫ 公開特許公報(A)

昭63-189933

⑤ Int. Cl.⁴

G 06 F 3/16
15/20

識別記号

3 4 0
3 0 1

庁内整理番号

P-7341-5B
T-7218-5B

③ 公開 昭和63年(1988)8月5日

審査請求 未請求 発明の数 1 (全4頁)

⑬ 発明の名称 文章読み上げ装置

⑩ 特 願 昭62-22909

⑪ 出 願 昭62(1987)2月2日

⑭ 発 明 者 内 山 昌 彦 神奈川県川崎市中原区上小田中1015番地 富士通株式会社
内

⑯ 出 願 人 富士通株式会社 神奈川県川崎市中原区上小田中1015番地

⑰ 代 理 人 弁理士 井 柁 貞一

明 細 書

1. 発明の名称

文章読み上げ装置

2. 特許請求の範囲

文章解析部(2)が入力される文章データを解析して、該文章データの単語を予め登録された単語の読みと照合してそれぞれの単語に読みを出力し、音声合成部(3)で音声合成して合成音で読み上げる装置において、

前記文章解析部(21,22)に接続された未知語読み変換部(5)と該未知語読み変換部(5)に接続された漢字辞書部(6)とを設け、前記前記文章解析部(21,22)に登録されていない前記文章データの単語は前記未知語読み変換部(5)に出力して処理され、前記登録されていない単語で、1文字の漢字とかなで構成される単語の漢字は訓練読みで読みを出力し、複数の漢字で構成される単語の漢字は音読みで読みを出力して音声合成部(3)に輸入されることを特徴とする文章読み上げ装置。

3. 発明の詳細な説明

(概要)

本発明は文章読み上げ装置に入力される文章データに文章解析部に登録されていない単語があった場合、その単語に与える読みに関するものであって、未知語読み変換部と漢字辞書部とを設け、1漢字とかなで構成される単語の漢字は訓練読みにし、複数の漢字は音読みにすることによって未登録単語に読みを付与して、読み上げができるようにする。

(産業上の利用分野)

本発明は文章読み上げ装置に関するものであって、特に、入力される文章データに文章解析部に登録されていない単語があった場合、その単語に与える読みに関するものである。

文字コードで構成された文章データを合成音声によって読み上げる文章読み上げ装置は、新聞社等における原稿とその原稿から作成された文章データとの読み合わせ校正に利用される。

即ち、文章読み上げ装置が文章データを電子的に合成された合成音声で読みあげ、校正者は原稿を見て校正する。

この場合、文章データには文章解析部に登録されていない単語、例えば造語、新語等がしばしば含まれており、その都度文章読み上げ装置がとまってしまうとその部分の校正確認が煩わしく、校正作業工数が増加することになる。

従って、未登録語はなんらかの読みを与えて読むことができるような文章読み上げ装置が要望されていた。

(従来の技術)

第4図は従来の文章読み上げ装置の構成ブロック図を示す。

図において、文字コードで構成された文章データがファイル装置、あるいは回線伝送によって文章入力部1に入力する。

文章解析部2は文章入力部1から出力された文章データを単語に区分し、内蔵する単語辞書と照

合して、区分された単語に読み、アクセント、発声高低等の単語発音に必要な要素、また文章として読み上げる場合の単語発音の抑揚等を付加して音声合成部3に出力する。

音声合成部3は、これら発音要素のデジタル信号を音声のアナログ信号に変換し、スピーカ4から出力される。

このような文章読み上げ装置で文章解析部2の単語辞書に登録されていない単語、即ち、未知語があって、発音要素に解析できない場合は、音声合成ができず、文章の読み上げはその部分で停止される。

そして、その場合は文字コードを印刷、あるいはディスプレイ表示して、原稿との正誤を確認する。

(発明が解決しようとする問題点)

この従来の方式では、例えば新聞社で使用されるような文章では、新語、造語、略語等が頻りに現れるため、未知語は多くなり、文章の読み上げ

はその都度中断される。

未知語によって文章の読み上げが中断され、その確認する手段によって、原稿校正作業は著しく阻害される。

本発明はこのような点に鑑みて創作されたものであって、未知語に対してなんらかの読みを与え、文章読み上げを中断しない文章読み上げ装置を提供することを目的としている。

(問題点を解決するための手段)

上記した目的を達成するために、文章読み上げ装置に未知語読み変換部と漢字辞書部とを設け、文章解析部に接続する。

そして、文章解析部に登録されていない未知語について、1文字の漢字とかなとで構成される単語の漢字は、漢字辞書部に照合して訓読みで読みを付与し、複数の漢字で構成される未知語の漢字は、音読みで読みを付与するようにする。

(作用)

文章解析部で解析できない未知語は、未知語読み変換部に出力され、未知語読み変換部は漢字辞書部に登録された漢字と照合して、未知語の単語構成に応じて漢字に訓読み、あるいは音読みを与え、音声合成部に出力する。

この未知語の処理によって、文章読み上げ装置は未知語によって中断されることがなくなる。

しかも、文章解析部における単語辞書にすべての単語の登録を期待することなく、登録される単語を常用単語に限り、未登録単語は未知語として未知語読み変換部の読み任せにすることができる。

こうすることによって、単語辞書の登録数を減らすと文章解析文章の単語照合の平均時間が減少し、文章読み上げが円滑に行い得ることになる。

(実施例)

第1図は本発明の文章読み上げ装置の一実施例の構成ブロック図を示す。

なお、全図を通じて同一符号は同一対象物を示

す。

文章解析部21は単語辞書に登録されていない単語を未知語として弁別し、未知語読み変換部5に出力する機能が付与される。

従って、入力された文章データの未知語は、文章解析部を経て、未知語読み変換部5に出力される。

未知語読み変換部5は、漢字辞書部6に読みを照合して未知語に読みをつけ、音声合成部3に出力する。

漢字辞書部6は、漢字読みの第1の登録方式として、1つの漢字に音読みと訓読みとを1つずつをもっていて、未知語が漢字とかなで構成されていると、その漢字に訓読みを抽出し、未知語が複数の漢字で構成されている場合は、その漢字に音読みを付与する。

この方式の場合、必ずしも正しい未知語の読みを期待することなく、ただ文章読み上げ時の読みが対応する1つの漢字に特定されればよいとするものである。

このようにすることによって、未知語には文章解析文章22で使用された最新の使用読みが付与される。

未知語が例えば「行革審」と云う複数の漢字からなる単語であれば、これら漢字が読み複数の音読みをもつ場合でも、文章データの始めに「行政改革審議会」という単語があって「ギョウセイカイカクシンギカイ」と読まれているような場合に、「行革審」は未知語とされても「ギョウカクシン」と読まれることになる。

なお、未知語の読みは文章解析部の単語辞書に登録して、次の文章データの単語入力に対処するようすることも容易に可能である。

(発明の効果)

以上述べてきたように、本発明によれば、文章読み上げ装置の文章解析部に登録されていない未知語であっても読みが付与され、中断することなく読み上げを続けることができ、実用的には極めて有用である。

第2図は本発明の文章読み上げ装置の他の実施例の構成ブロック図、

第3図は漢字読みの第2の登録方式を説明する図を示す。

文章解析部22は、未知語読み変換部5に未知語を出力するとともに、漢字辞書部6の漢字の読みの標識欄cに標識を付ける機能を有する。

第2の登録方式として、漢字とが対応している文章解析部22の漢字読みを第3図のように、漢字毎に、漢字aに対する読み欄b、各読みに対する標識欄cを有する。

例えば「行」であれば、「オコナ(う)〔訓〕、ギョウ(音)、コウ(音)等の読みを記憶し、「政」では、「マツリゴト(訓)、セイ(音)」を記憶している。

そして、標識欄cは文章解析文章22が漢字に読みを付与するときに使用した音読みに標識をつけ換える。

未知語読み変換部5は、未知語の漢字読みには漢字辞書部6の標識のある読みを使用する。

4. 図面の簡単な説明

第1図は、本発明の文章読み上げ装置の一実施例の構成ブロック図、

第2図は、本発明の他の実施例の構成ブロック図、

第3図は、第2図の漢字辞書を説明する図、

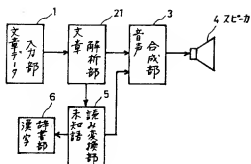
第4図は、従来例の構成ブロック図である。

図において、

- 1は文章データ入力部、
- 2, 21, 22は文章解析部、
- 3は音声合成部、
- 4はスピーカ、
- 5は未知語読み変換部、
- 6は漢字辞書部である。

代理人 弁理士 井桁貞一





本発明の一実施例の構成ブロック図

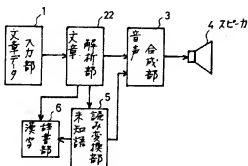
第 1 図

a	b	c
行	オコナ(ウ)	
	ギョウ	○
	コウ	

政	マツリゴト	
	セイ	○

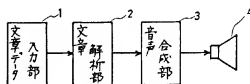
漢字辞書と説明表図

第 3 図



本発明の他の実施例の構成ブロック図

第 2 図



従来例の構成ブロック図

第 4 図